Biophysical Journal, Volume *110*

Supplemental Information

Conformational and Thermodynamic Landscape of GPCR Activation from Theory and Computation

Sijia S. Dong, William A. Goddard, III, and Ravinder Abrol

SUPPORTING MATERIAL

Conformational and Thermodynamic Landscape of GPCR Activation From Theory and Computation

Sijia S. Dong, William A. Goddard III*, Ravinder Abrol[1,]*

Materials and Process Simulation Center (MSC),

California Institute of Technology, Pasadena, CA 91125, United States

Corresponding authors: wag@wag.caltech.edu; abrolr@csmc.edu

[1] Current address: Departments of Biomedical Sciences and Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, United States

I. Additional Computational Details

1. Pre-processing of crystal structure templates for structure prediction

For each validation case, hβ_2AR and hM2, we determined the range of TM regions by taking a consensus of the helix assignment in the Protein Data Bank (PDB) of its active state structure and inactive structure, taken from the Orientations of Proteins in Membranes (OPM) database.(1) We then cropped out the TM regions and added missing atoms using tleap in AmberTools1.4.(2) To have meaningful comparisons of the energetics of structures sampled from inactive state crystal structure and from active state crystal structure, we minimized all these structures (TM regions only) using MPSim(3) with a convergence criterion of root mean square force (RMS force) = 0.25 kcal mol^{-1} Å^{-1} before proceeding to the conformational sampling steps.

2. Constructing the hybrid templates

For Validation Method 2.1

We aligned the active state crystal structure to the inactive-state crystal structure using Visual Molecular Dynamics (VMD).(4) We considered only the backbone atoms in TM1-5 and TM7 in the alignment. Then we replaced the TM6 in the inactive-state crystal structure with the TM6 in the active-state crystal structure.

For Validation Method 3.3

We obtained an active state candidate, denoted S1, from coarse SuperBiHelix sampling of the inactive-state crystal structure. Then we aligned the active-state crystal structure to S1 using VMD. We considered only the backbone atoms in the alignment. At last, we replaced the TM6 in S1 with the TM6 in the active-state crystal structure.

For application to hSSTR5

We obtained a homology model of the target protein using an inactive-state crystal structure as the template, and did coarse SuperBiHelix to obtain an active-state candidate T1. We also obtained the homology model of the target protein using an active-state crystal structure as the template, and name this model as T2. Then we aligned T2 to T1 using VMD. We considered only the backbone atoms in the alignment. At last, we replaced the TM6 in T1 with the TM6 in T2. The x,y coordinates of the hydrophobic center (HPC) of the hybrid template remain to be that of T1.

3. Energy scoring function in structure prediction

Whenever energy ranking was performed in structure prediction procedures, we used energy E_{CNti} which is the average energy of four types of energy: a) the total energy of the charged protein (CTotal), b) the interhelical energy of the charged protein (CInterH), which neglects the intrahelical energy of each chain, c) the total energy (NTotal), and d) the interhelical energy (NInterH) of the neutralized protein.

4. Ligand docking

Scanning regions in protein for docking

The first step of GenDock(5, 6) is DarwinDock, which modifies the protein structure to replace the six types of hydrophobic residues by alanine, and then samples the complete set of poses for regions that could potentially bind a ligand. To do this sampling, the potential binding region is filled by SphGen with "spheres" having 2 Å overlaps with each other and the spheres classified into "boxes" of 10 Å sides. Boxes containing 75 or more spheres were kept. For docking purpose, we have discarded all spheres except for those that are in the extracellular half of the GPCR TMDs and are not potentially in contact with the membrane lipids (i.e. are in the interior of the GPCR helix bundle).

GenDock

For each ligand conformation and for the "spheres" selected in the previous step, we generated 200 000 poses without energy evaluation aiming at providing a complete set of poses. The poses were clustered into ~8000-9000 Voronoi families based on RMSD and the binding energy of the family head evaluated. Then for the top 10% of families, we evaluated the energy for all children. Then we selected the top 50 based on each three energy scores: polar energy, hydrophobic energy, and total energy. Then for these 150, we dealanized (mutating alanine back to the original hydrophobic residues) and optimized the side chains using SCREAM. Then the protein-ligand complexes (poses) were subject to minimization for 50 steps.

Docking procedure of agonists to hSSTR5

For each ligand, we first did a conformational search using MacroModel 9.7(7) in Maestro 9.1.(8) For L-817,818, we sampled 6 rotatable bonds . For F21, the –CH2-CH2-Ph group was first replaced by a methyl group. This modified molecule is labeled F21m. For F21m, 7 rotatable bonds were sampled. The conformational search was a torsional sampling of the rotatable bonds that could cause large conformational changes using a Monte Carlo Multiple Minimum (MCMM) method(9, 10) with the force field OPLS 2005.(11) The energy window for the generated structure to be kept was set to be 10.04 kcal mol^{-1}. Then we clustered the resulting conformations with an RMSD cutoff of 2 Å. For L-817,818 and F21m, the clustering resulted in 18 and 8 distinct conformations respectively. Then each F21m conformation was modified into F21 by adding back the –CH2-Ph group with each of the 9 possible conformations. The ligand conformations were then docked to each candidate protein structure using our standard docking strategy, GenDock,(5, 6) described above. The charge distribution used in docking was obtained by the Mulliken population analysis using B3LYP/6-311G** in Jaguar 7.6.(12) For each ligand and each docked protein, we collected 1000 lowest unified-cavity (UCav) energy complexes. Then for each ligand, we collected 15 lowest UCav energy complexes and matched ligands at different positions in these complexes into each of the protein conformations in these complexes. We did a simulated annealing on the resulting complexes' ligands and residues within 10 Å from the ligands. In the end, we minimized each of the complexes. The final complexes were scored by snap binding energy (SnapBE). For each ligand, the lowest energy complex among all complexes with an active-state GPCR was selected as the final active-state pose, and the lowest energy complex among all complexes with an inactive-state GPCR was selected as the final inactive-state pose.

5. Molecular dynamics (MD) simulation

Preparing the crystal structures of hβ$_2$AR and the G$_s$ protein for MD simulation

Starting from the active-state crystal structure coupled with the G_s protein (PDB ID: 3SN6), we modeled the missing loops in the Gα domain of the G protein using the SWISS-MODEL server,(13) and built the missing loops in the GPCR with a Monte Carlo technique that grows geometrically allowed loop structures from the two TM ends. We then relaxed the modeled loops by simulated annealing of 10 cycles (lowest temperature: 50 K, highest temperature: 600 K) followed by minimization till RMS force reached 0.25 kcal mol^{-1} Å^{-1}. For the inactive-state case, we aligned the inactive-state crystal structure to the active-state crystal structure to match the agonist into the inactive-state structure. We then minimized the complex till RMS force reached 0.25 kcal mol^{-1} Å^{-1} and did simulated annealing of the binding site within 6 Å of the ligand, followed by minimization.

Loop building of hSSTR5

We modeled ICL1, ECL1 and ECL2 from mOPRM crystal structure (PDB ID: 4DKL)(14) using homology modeling. The remaining loops, ICL2, ICL3 and ECL3, were built with a Monte Carlo technique that grows geometrically allowed loop structures from the two fixed TM ends. Then we added the C-terminus of hSSTR5 up to the C-terminus of Helix8 (C320) by attaching Helix8 of the template after aligning their NPxxY motifs followed by mutating to hSSTR5. In addition, we added the N-terminus from residue 36 to 38. Minimization was then carried out on the final structure while keeping the TM domains (except the end residues) fixed until energy was converged.

Modeling Gα$_i$

We used SWISS-MODEL(13) to model Gα$_i$ (UniProt ID: P63096, canonical) from Gα$_s$ in the G protein heterotrimer crystalized with hβ$_2$AR (PDB ID: 3SN6).(15) Using MPSim,(3) we minimized Gα$_i$ in vacuum until the RMS force is lowered to 0.1 kcal mol^{-1} Å^{-1}. Then Gα$_i$ was placed with ActiveConf2 by aligning ActiveConf2 to hβ$_2$AR in 3SN6 and aligning Gα$_i$ to Gα$_s$ in 3SN6. Only backbone atoms were considered in the alignment.

For the ActiveConf2+Gα$_i$ complex, we used MPSim to optimize all loops plus one more residue at each end of each TM the GPCR together with residues 352 to 354 of Gα$_i$ in the ActiveConf2+Gα$_i$ complex.

Building the lipid/water environment for hβ_2AR

1) For the agonist+GPCR complex: We used Visual Molecular Dynamics (VMD),(4) to insert the prepared complex into a 75 Å × 85 Å lipid bilayer structure (for the inactive state) or into a 75 Å × 95 Å lipid bilayer structure (for the active state). This system was then placed into a water box with a total of ~11300 (for the inactive state) water molecules or ~13000 (for the active state) water molecules in the 15 Å and 35 Å thick space on the extracellular and intracellular sides of the lipid bilayer.

2) For the agonist+GPCR+Gα_s complex: Using VMD, the prepared agonist+GPCR+Gα_s complex was inserted into a 120 Å × 130 Å lipid bilayer structure. This system was then placed into a water box with a total of ~39500 water molecules in the 15 Å and 60 Å thick space on the extracellular and intracellular sides of the lipid bilayer.

3) For Gα_s alone: Using VMD, the prepared Gα_s protein was placed into a rectangular water box with 10 Å thick of water padded on each of x-, y-, z- direction of the protein. There were a total of ~19400 water molecules.

4) For the agonist alone: Using VMD, the agonist taken from the prepared agonist+GPCR+Gα_s complex was placed into a rectangular water box with 15 Å thick of water padded on each of x-, y-, z- direction of the molecule. There were a total of ~1700 water molecules.

For all the three cases above, in the end, Na^+ and Cl^- ions were placed into the system using tleap for a physiological NaCl concentration (0.9% w/v) and a neutral system.

Building the lipid/water environment for hSSTR5

1) For ligand-bound GPCR: Using VMD, the final hSSTR5 structure with loops built was inserted into a 75 Å × 75 Å lipid bilayer structure. The system was then placed into a water box with a total of ~8300 water molecules in the 15 Å and 25 Å thick space on the extracellular and intracellular sides of the lipid bilayer.

2) For agonist+GPCR+Gα_i complex: Using VMD, the agonist+GPCR+Gα_i complex was inserted into a 120 Å × 100 Å lipid bilayer structure. The system was then placed into a water box with a total of ~29600 water molecules in the 15 Å and 60 Å thick space on the extracellular and intracellular sides of the lipid bilayer.

3) For $G\alpha_i$ alone: Using VMD, the modeled $G\alpha_i$ was placed into a rectangular water box with 10 Å thick of water padded on each of x-, y-, z- direction of the protein. There were a total of ~20800 water molecules.
4) For the agonist alone: Using VMD, the agonist taken from the agonist+GPCR+$G\alpha_i$ complex was placed into a rectangular water box with 15 Å thick of water padded on each of x-, y-, z- direction of the molecule. There were a total of 1855 water molecules.

For all the three cases above, in the end, Na^+ and Cl^- ions were placed into the system with tleap for a physiological NaCl concentration and a neutral system.

MD simulation protocol

MD_Step1) With the ligand and TM regions of the receptor fixed, the loops and Helix8 of the receptor, lipids and water molecules in the system were minimized for 10000 steps using the conjugate gradient method. In the case of the agonist+GPCR+Gα complex, Gα is fixed in this step too. In the case of Gα alone in solvent, only Gα is fixed in this step, and the minimization was carried out for 5000 steps.
MD_Step2) With the ligand and TM regions of the receptor fixed, the loops and Helix8 of the receptor, the lipids and water molecules were equilibrated at 310 K and 1 atm for 1 ns using the NPT ensemble. This allowed the water molecules to defuse into the ligand-protein system filling any artificial voids in the simulation system. In the case of the agonist+GPCR+Gα complex, Gα is fixed in this step too. In the case of Gα alone in solvent, only Gα is fixed in this step, and the equilibration was for 1.5 ns.
MD_Step3) The whole system was then minimized for 10000 steps using the conjugate gradient method. In the case of Gα alone in solvent, the minimization was carried out for 5000 steps.
MD_Step4) The whole system was heated from 0 K to 310 K in hundreds of ps and then equilibrated using the NPT ensemble for a total of 51 ns MD simulation.

For the simulation of apo-GPCR+Gα, we took the last frame of agonist+GPCR+Gα complex from the above procedure, removed the ligand from the system, adjusted the number of Na^+ or Cl^- ions to make the system neutral again if the ligand was charged, and repeated the above MD protocol MD_Step1) to MD_Step4) with the all parts of the proteins fixed in MD_Step1).

II. Supporting Tables and Figures

TABLE S1 Comparison of R_{36} between inactive-state and active-state GPCR crystal structures. R_{36} is the minimal approach distance between the backbone atoms of the intracellular ends of TM3 and TM6, defined in the Materials and Methods section. R_{36} of active state structure is $R_{36}^{(a)}$; R_{36} of inactive state structure is $R_{36}^{(i)}$.

Protein (PDB ID)	Protein Description	Coupled Protein	R_{36} (Å)	$R_{36}^{(a)} - R_{36}^{(i)}$ (Å)
hβ_2AR (3SN6)(15)	active	G protein heterotrimer	13.83	4.12
hβ_2AR (2RH1)(16)	inactive	-	9.71	-
bRho (3PQR)(17)	active (metarhodopsin II)	Gα C-terminal fragment	11.44	3.69
bRho (1U19)(18)	inactive (rhodopsin)	-	7.75	-
hM2 (4MQS)(19)	active	G protein mimetic camelid antibody fragment	11.34	2.84
hM2 (3UON)(20)	inactive	-	8.50	-
mOPRM (5C1M)(21)	active	G protein mimetic camelid antibody fragment	12.21	5.08
mOPRM (4DKL)(14)	inactive	-	7.13	
hAA_{2A}R (2YDO)(22)	agonist-bound	-	7.48	1.34
hAA_{2A}R (3EML)(23)	antagonist-bound	-	6.14	-

TABLE S2 Summary of structural features and energies of hβ_2AR and hM2 optimal structures generated from different methods. The "Structure Identifier" corresponds to the numbers in Fig. 1. Each RMSD value is between the backbone atoms of the resulting optimal structure of the particular method and the backbone atoms of the active-state crystal structure preprocessed according to Supporting Materials and Methods. For hM2, the second best choice of 3.2 is also listed for comparison. Unlike hβ_2AR for which the inactive state is 73.8 kcal mol^{-1} more stable than the active state, this value for hM2 is only 3.7 kcal mol^{-1}. Therefore, the number of 7-helix bundles built from SuperBiHelix results is increased to 2500 to capture more candidates in the active-like regime. (For hβ_2AR, building 1000 or 2500 bundles has the same the final results shown in this table.) While hM2's lowest-energy structure from AfromI_Fine that satisfies the active-like R_{36} criterion (upper row of 3.2 in the table) has a much larger RMSD than Structure 3.1 to the active-state crystal structure, the second-lowest-energy structure (lower row of 3.2 in the table) has an improved RMSD comparing to 3.1. This suggests selecting a small number of diverse structures from the potential energy well may help in active-state structure prediction as well.

Structure / Method Identifier	Method Name	Method Description			$h\beta_2AR$			hM2		
		Sampling Target	Starting Structure	Sampling Method	E_{CNti} (kcal mol^{-1})	RMSD to Target Crystal Structure (Å)	R_{36} (Å)	E_{CNti} (kcal mol^{-1})	RMSD to Target Crystal Structure (Å)	R_{36} (Å)
Active crystal structure	-	Active State	Active crystal structure	None	-	0.00	13.72	-	0.00	11.14
Inactive crystal structure	-	Active State	Active crystal structure	None	-	2.48	9.71	-	2.30	8.55
1.1	AfromA_Coarse	Active state	Active crystal structure	Coarse	-146.8	0.04	13.68	-168.5	0.37	11.15
1.2	AfromA_Fine	Active state	Active crystal structure	Fine after Coarse	-146.8	0.04	13.68	-168.5	0.37	11.15
2.1	AfromH_Coarse	Active state	Active crystal TM6 + inactive other TMs	Coarse	-158.0	1.41	15.12	-96.9	1.55	13.60
2.2	AfromH_Fine	Active state	Active crystal TM6 + inactive other TMs	Fine after Coarse	-196.9	1.48	14.64	-100.3	1.69	13.45
3.1	AfromI_Coarse	Active state	Inactive crystal structure	Coarse	-95.0	1.79	12.85	-77.4	1.94	14.45
3.2	AfromI_Fine	Active state	Inactive crystal structure	Fine after Coarse	-114.5	1.86	13.41	-127.2	2.68	15.20
								-111.6	1.92	15.22
3.3	AfromIH_Fine	Active state	Inactive crystal structure	Fine after replacing TM6 of 3.1 result by active crystal structure TM6	-169.5	1.47	14.03	-113.2	1.52	14.44
Inactive crystal structure	-	Inactive state	Inactive crystal structure	None	-	0.00	9.71	-	0.00	8.55
4.1	IfromI_Coarse	Inactive state	Inactive crystal structure	Coarse	-220.6	0.07	9.72	-168.8	0.04	8.54
4.2	IfromI_Fine	Inactive state	Inactive crystal structure	Fine after Coarse	-220.6	0.07	9.72	-168.8	0.04	8.54

TABLE S3 Summary of structural features and energies of a) hβ_2AR and b) hM2 optimal structures generated from different methods. The "Method Identifier" corresponds to the numbers in Fig. 1. Each RMSD value is between the backbone atoms of the resulting optimal structure of the particular method and the backbone atoms of the active-state crystal structure. For TM4, none-zero Δφ values do not have as big effect as Δφ on the orientations of other TMs because the absolute value of θ for TM4 is close to zero.

a)

Structure / Method Identifier	Method Name	Δθ (°)	Δφ (°)							Δη (°)							E_{CNti} (kcal mol^{-1})	RMSD to crystal active (Å)	R_{36} (Å)
		H1 – H7	H1	H2	H3	H4	H5	H6	H7	H1	H2	H3	H4	H5	H6	H7			
Active crystal structure	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	0.00	13.72
Inactive crystal structure	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	2.48	9.71
1.1	AfromA_Coarse	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-146.8	0.04	13.68
1.2	AfromA_Fine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-146.8	0.04	13.68
2.1	AfromH_Coarse	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-158.0	1.44	15.12
2.2	AfromH_Fine	0	0	0	0	-15	0	15	15	0	0	0	0	0	0	0	-196.9	1.48	14.64
3.1	AfromI_Coarse	0	0	0	0	0	0	-45	0	0	0	0	0	0	-30	0	-95.0	1.79	12.85
3.2	AfromI_Fine	0	0	0	0	-15	0	-75	30	0	0	0	0	0	-60	0	-114.5	1.86	13.41
3.3	AfromIH_Fine	0	0	0	0	15	0	15	15	0	0	0	0	0	0	0	-169.5	1.47	14.03
4.1	IfromI_Coarse	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-220.6	2.47	9.72
4.2	IfromI_Fine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-220.6	2.47	9.72

b)

Structure / Method Identifier	Method Name	Δθ (°)							Δφ (°)							Δη (°)							E_{CNti} (kcal mol^{-1})	RMSD to crystal active (Å)	R_{36} (Å)
		H1	H2	H3	H4	H5	H6	H7	H1	H2	H3	H4	H5	H6	H7	H1	H2	H3	H4	H5	H6	H7			
Active crystal structure	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	0.00	11.14
Inactive crystal structure	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	2.30	8.55
1.1	AfromA_Coarse	0	0	0	0	0	0	0	0	0	0	-45	0	0	0	0	0	0	0	0	0	0	-168.5	0.37	11.15
1.2	AfromA_Fine	0	0	0	0	0	0	0	0	0	0	-45	0	0	0	0	0	0	0	0	0	0	-168.5	0.37	11.15
2.1	AfromH_Coarse	0	0	0	0	0	0	-15	0	0	0	0	0	0	-90	0	0	0	0	0	-30	0	-96.9	1.55	13.60
2.2	AfromH_Fine	0	0	0	0	0	0	0	0	0	0	-15	15	15	-60	0	0	0	0	0	-30	0	-100.3	1.69	13.45
3.1	AfromI_Coarse	0	0	0	0	15	15	-15	0	0	0	0	0	-90	0	0	0	0	0	0	0	0	-77.4	1.94	14.45
3.2	AfromI_Fine	0	0	0	0	30	15	0	0	0	0	-15	15	-60	0	0	0	0	0	-30	0	0	-127.2	2.68	15.20
		0	0	0	0	15	15	0	0	0	0	-15	30	-60	15	0	0	0	0	0	0	0	-111.6	1.92	15.22
3.3	AfromIH_Fine	0	0	0	0	0	0	0	0	0	0	0	15	0	15	0	0	0	0	0	-30	0	-113.2	1.52	14.44
4.1	IfromI_Coarse	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-168.8	0.04	8.54
4.2	IfromI_Fine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-168.8	0.04	8.54

TABLE S4 Summary of starting structures (Methods 3.1, 4.1) and best resulting structures (Method 3.2, 3.3, 4.2.x) of sampling hSSTR5 conformations using different strategies. Method 3.x are active state samplings, and Method 4.x are inactive state samplings. The original homology template is from the mOPRM inactive-state crystal structure. The methods used are outlined in Fig. S1 and detailed in our previous publication.(25)

Structure / Method Identifier	Method Name	Δθ (°)		Δφ (°)							Δη (°)							E_{CNti} (kcal mol^{-1})	R_{36} (Å)
		H1 – H5, H7	H6	H1	H2	H3	H4	H5	H6	H7	H1	H2	H3	H4	H5	H6	H7		
Homology w/ mOPRM inactive-state crystal structure	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N/A	6.04
3.1	AfromI_Coarse	0	0	0	0	0	0	0	-90	0	0	0	0	0	30	270	0	-265.8	11.34
3.2 (ActiveConf1)	AfromI_Fine	0	0	0	0	0	0	0	-60	-15	0	0	0	0	30	270	0	-313.2	10.71
3.3 (ActiveConf2)	AfromIH_Fine	0	-15	0	0	-15	0	-30	-60	0	0	0	0	0	30	240	0	-306.6	10.20
4.1	IfromI_Coarse	0	0	0	0	0	90	0	0	0	0	0	0	-30	30	0	0	-303.3	6.01
4.2.1 (InactiveConf1)	IfromI_Fine	0	0	0	0	0	105	0	-15	0	0	0	0	-30	0	0	0	-350.7	7.18
4.2.2 (InactiveConf2)	IfromI_Fine	0	0	0	0	-15	60	30	15	0	0	0	0	0	60	0	0	-345.7	7.59
4.2.3 (InactiveConf3)	IfromI_Fine	0	0	0	0	-15	60	-30	15	0	0	0	0	0	30	0	0	-328.9	7.56

TABLE S5 Best pose of L-817,818 and F21 docked to active structures determined by lowest snap binding energy (SnapBE).

Pose Name	Ligand	Protein Structure	SnapBE (kcal mol^{-1})
L_i2	L-817,818	InactiveConf2	-118.32
L_a2	L-817,818	ActiveConf2	-106.05
F_i2	F21	InactiveConf2	-99.41
F_a2	F21	ActiveConf2	-92.39

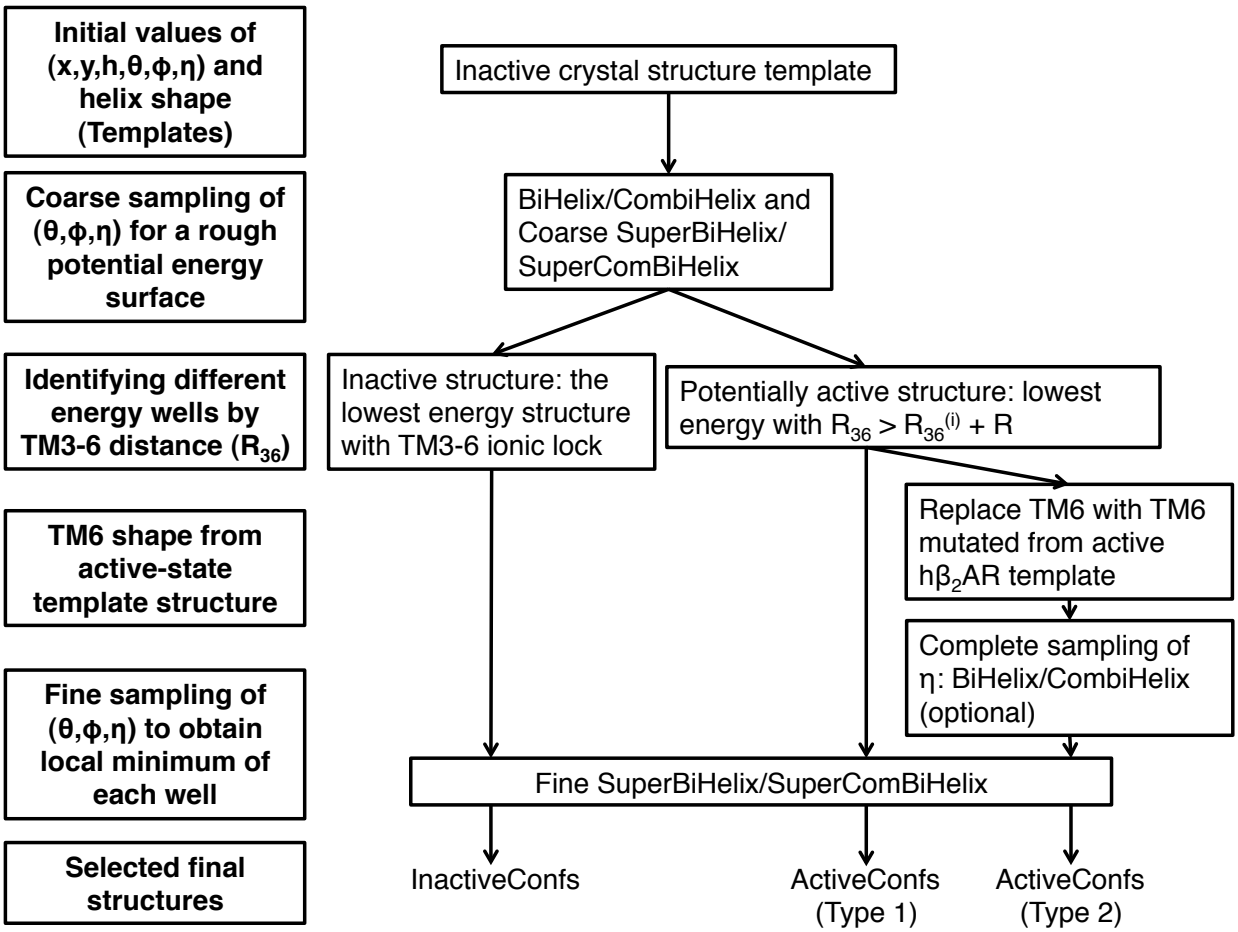

FIGURE S1 Schematic view of ActiveGEnSeMBLE. R_{36} is the minimal approach distance between the backbone atoms of the intracellular ends of TM3 and TM6, defined in the Materials and Methods section. $R_{36}^{(i)}$ is that of the inactive structure. R is a distance, usually chosen around 4 Å. Among all cases involved in this paper, the optional step was only carried out for hSSTR5 structure prediction, and it was found to be unnecessary as the final selected structures are the same as those without this step carried out. The sampling space of BiHelix is Δη from 0 to 360° in 30° increments. If not noted otherwise, the sampling space of Coarse SuperBiHelix is Δθ: 0, ±15°; Δφ: 0, ±45°, ±90°; Δη: 0, ±30°, selected angles from BiHelix/CombiHelix, starting with the best from BiHelix/CombiHelix. The sampling space of Fine SuperBiHelix is Δθ: 0, ±15°; Δφ: 0, ±15°, ±30°; Δη: 0, ±30°. The active state template for TM6 is chosen to be hβ_2AR (PDB ID: 3SN6) in this study because it is by far the only GPCR co-crystalized with a full G protein heterotrimer.

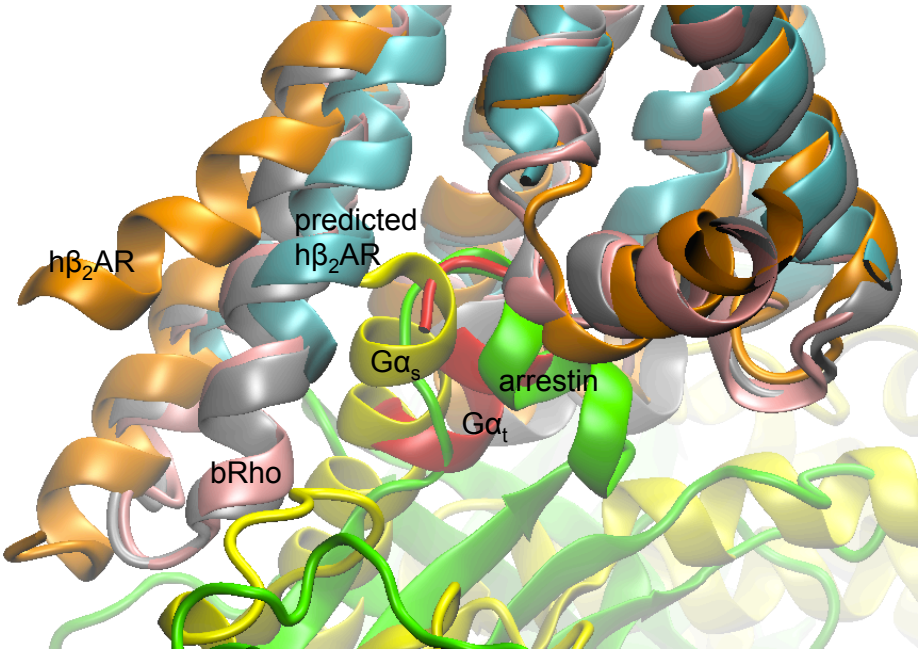

FIGURE S2 Predicted minimum-energy $h\beta_2AR$ structure (cyan) in one of the energy wells can accommodate arrestin (green) but not $G\alpha_s$ (yellow). An energy well is defined by clustering the structures with R_{36} values within 0.25 Å of each other. We positioned the $G\alpha_s$ protein and the arrestin by aligning $h\beta_2AR$ (orange) in its $G\alpha_s$-coupled crystal structure (PDB ID: 3SN6)(15), bRho (grey) in its arrestin-coupled crystal structure (PDB ID: 4ZWJ)(24) and bRho (pink) in its $G\alpha_t$-C-terminus-coupled crystal structure (PDB ID: 3PQR)(17) to the predicted $h\beta_2AR$ structure without changing the relative orientation within each crystal structure. We find the minimum energy structure of $h\beta_2AR$ in the $R_{36} \approx$ 12.4 Å well clashes with $G\alpha_s$ but not with arrestin. It is hard to determine by speculation whether it distinguishes arrestin with the G_i protein because the C-terminus helix of $G\alpha_t$ subunit (red) co-crystalized with the active-state bRho is similar in position with the part of arrestin inside the GPCR. (Therefore, similar analysis was not carried out on hM2 which couples with the G_i protein instead of G_s *in vivo*.) This structure is from the ensemble of structures generated using Method 3.2 which uses the inactive-state TM shapes for all TMs. It is characterized by $\Delta\theta = 0$ for all helices; $\Delta\varphi$ = -15°, 0, 0, 15°, 0, -30°, 15°; $\Delta\eta$ = 0, 0, 0, 0, 0, -30°, 0 with the values ordered from TM1 to TM7. It has E_{CNti} = -114.1 kcal mol^{-1}. The coordinates of this predicted structure are included in the Supporting Material as a separate file in the PDB format.

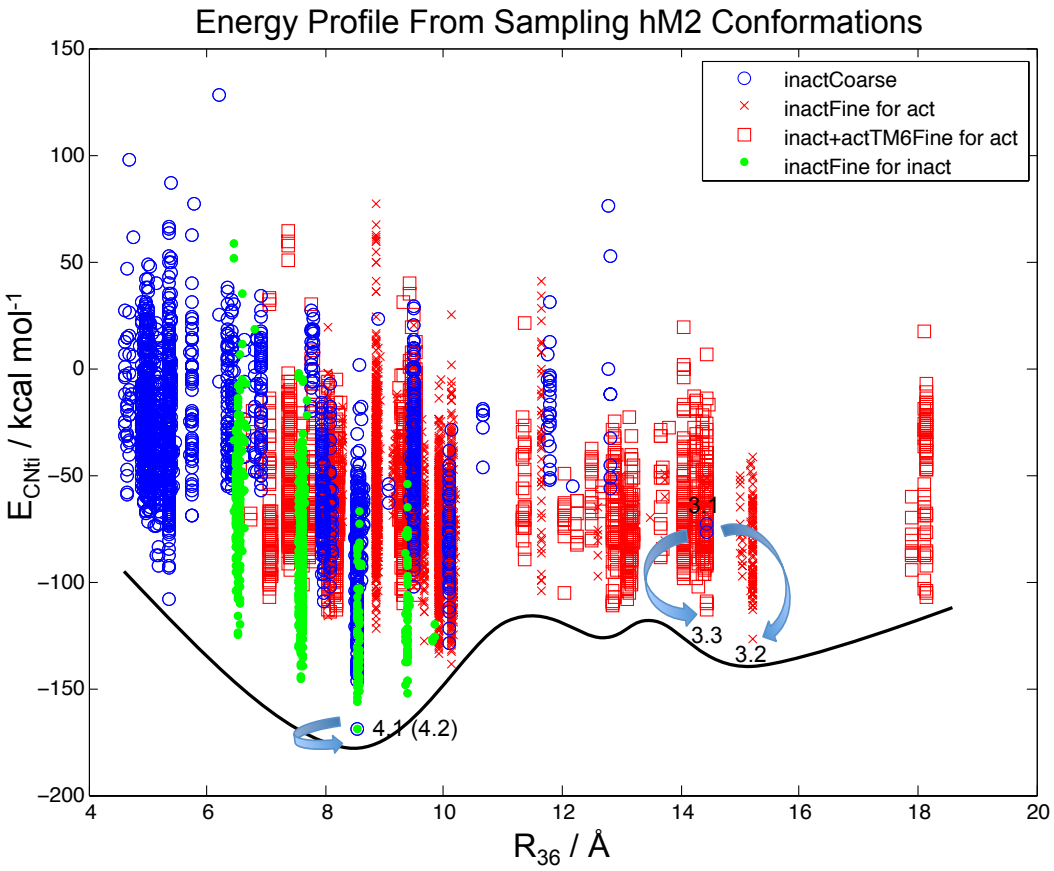

FIGURE S3 Potential energy profile from sampling hM2 conformations. The black curve is an illustration showing how our sampling results can qualitatively be translated into a potential energy curve using R_{36} as the x-axis and does not quantitatively represent any real data. Results of the coarse sampling starting from the inactive-state crystal structure are in blue circles. Starting from Structure 3.1, results of methods that generated Structures 3.2 and 3.3 are shown in red crosses and red squares respectively. Starting from the Structure 4.1, results of the fine sampling that generated Structure 4.2 are in green dots. Every blue arrow points from a starting structure towards the optimal structure from the corresponding sampling method.

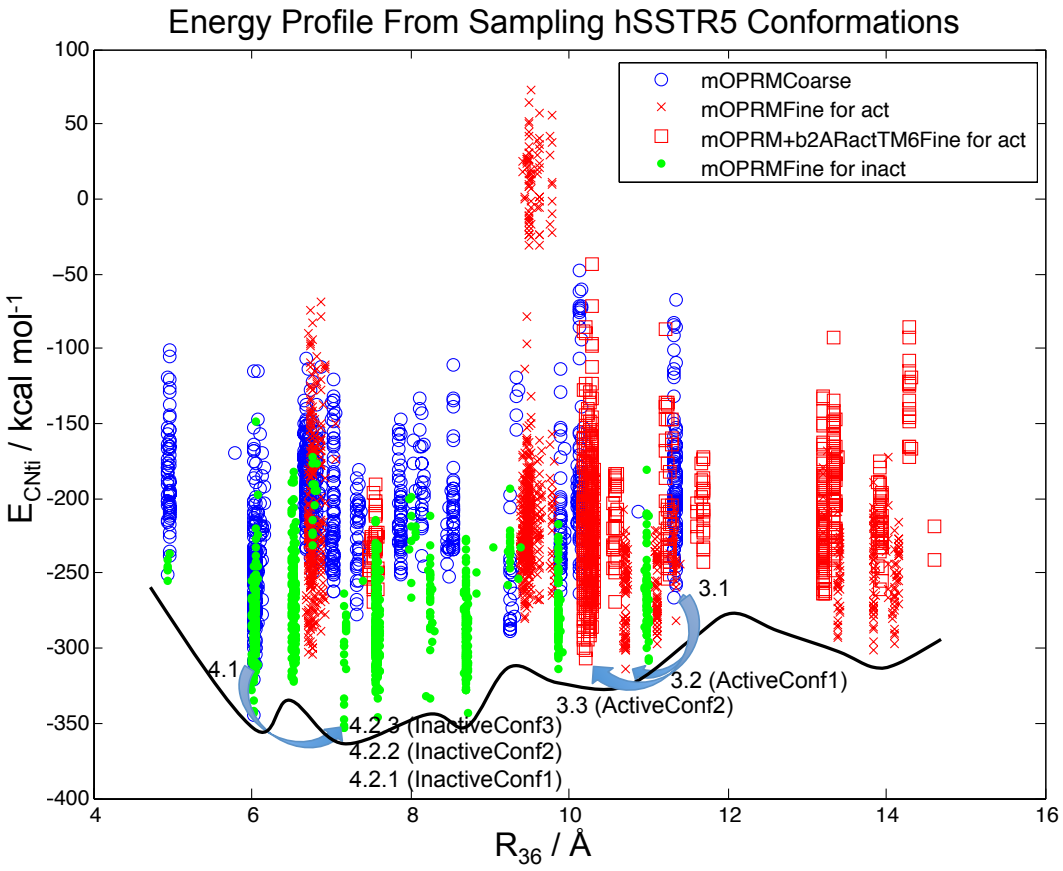

FIGURE S4 Energy profile from hSSTR5 samplings. The black curve is an illustration showing how our sampling results can qualitatively be translated into a potential energy curve using R_{36} as the x-axis and does not quantitatively represent any real data. Results of the coarse sampling are in blue circles. Starting from Structure 3.1, results of methods that generated Structures 3.2 (ActiveConf1) and 3.3 (ActiveConf2) are shown in red crosses and red squares respectively. Starting from the Structure 4.1, results of the fine sampling that generated Structures 4.2.1, 4.2.2 and 4.2.3 (InactiveConf1,2,3) are in green dots. Every blue arrow points from a starting structure towards the optimal structure from the corresponding sampling method.

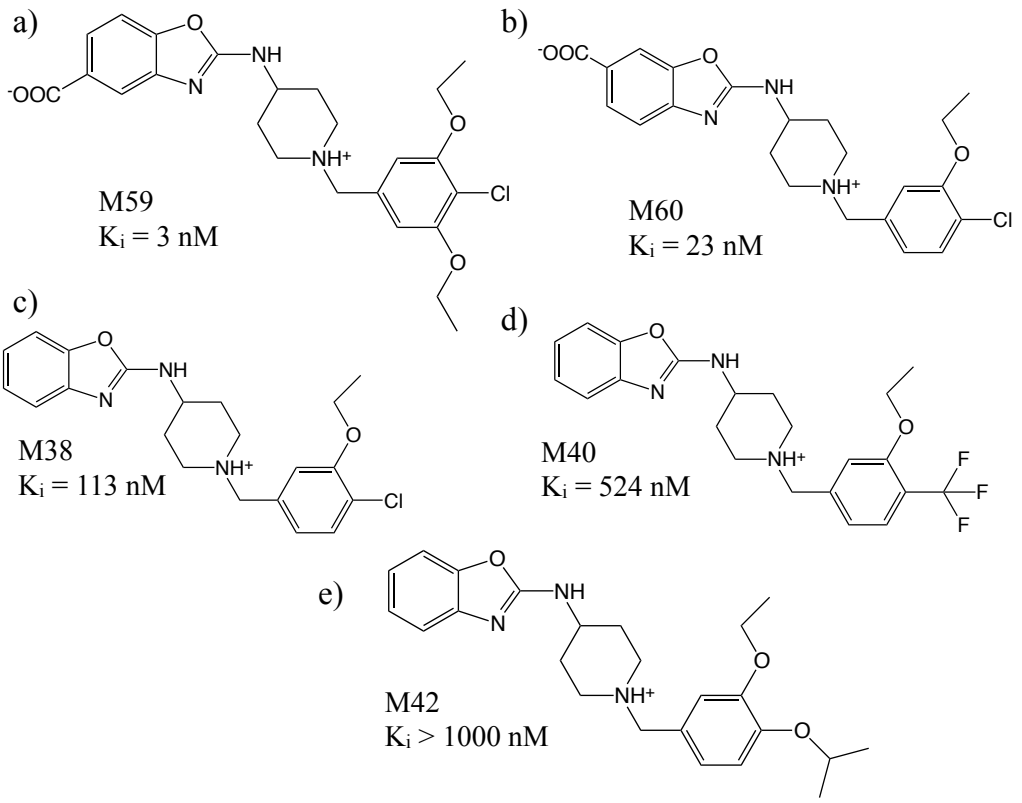

FIGURE S5 Antagonists series docked to hSSTR5 predicted structures.

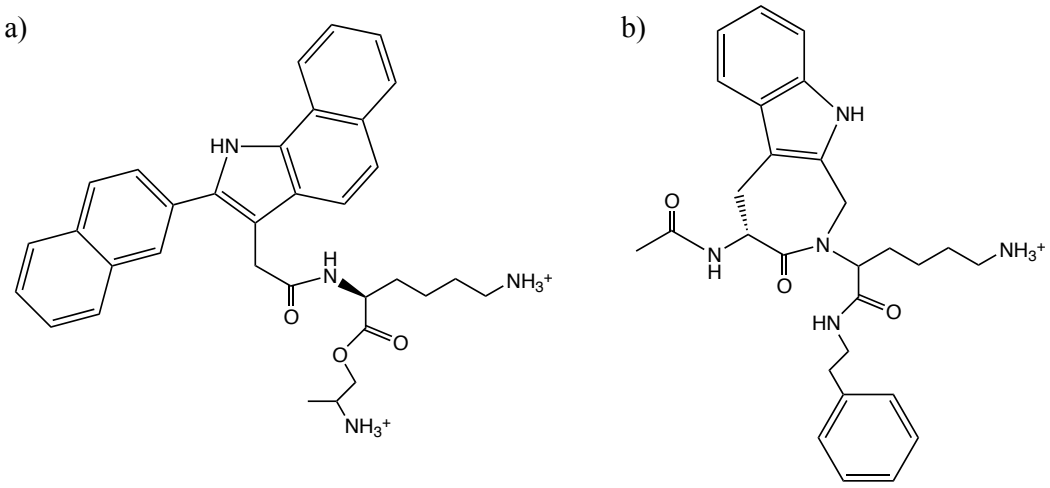

FIGURE S6 Structures of the docked agonists a) L-817,818 (K_i = 0.4 nM) and b) F21 (IC_{50} = 0.56 nM).

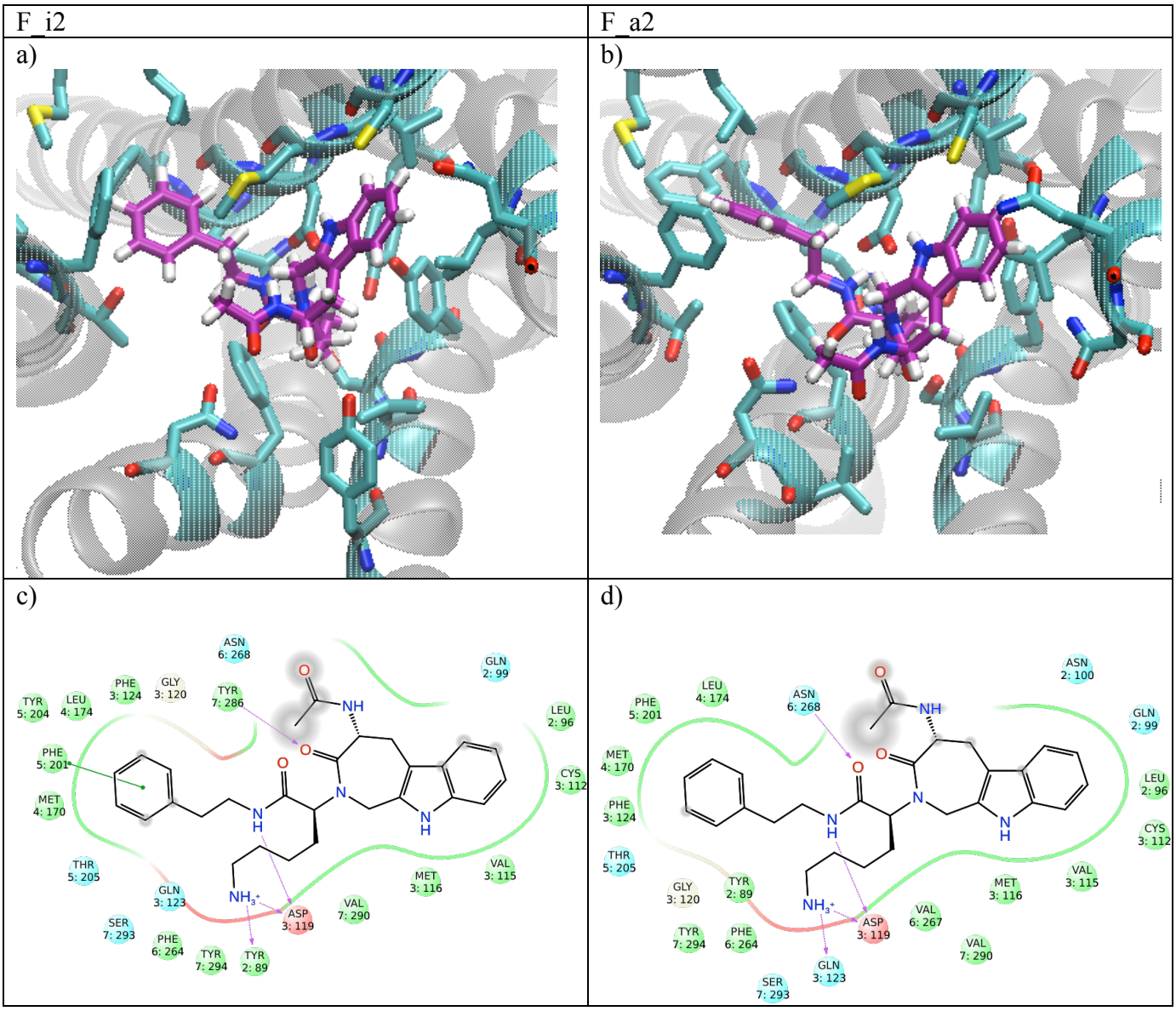

FIGURE S7 a) 3D visualization of the docking pose F_i2, the lowest energy complex with agonist F21. b) 3D visualization of the docking pose F_a2. The ligand is shown in purple. c) Ligand interaction diagram (LID) of the pose F_i2. d) LID of the pose F_a2. LIDs are generated by Maestro9.3.(26) The cutoff distance for the residues shown is 4.0 Å. Hydrophobic interaction: green; polar interaction: blue; hydrogen bonds (cutoff distance 2.5 Å): purple arrows; π-π stacking: straight green lines.

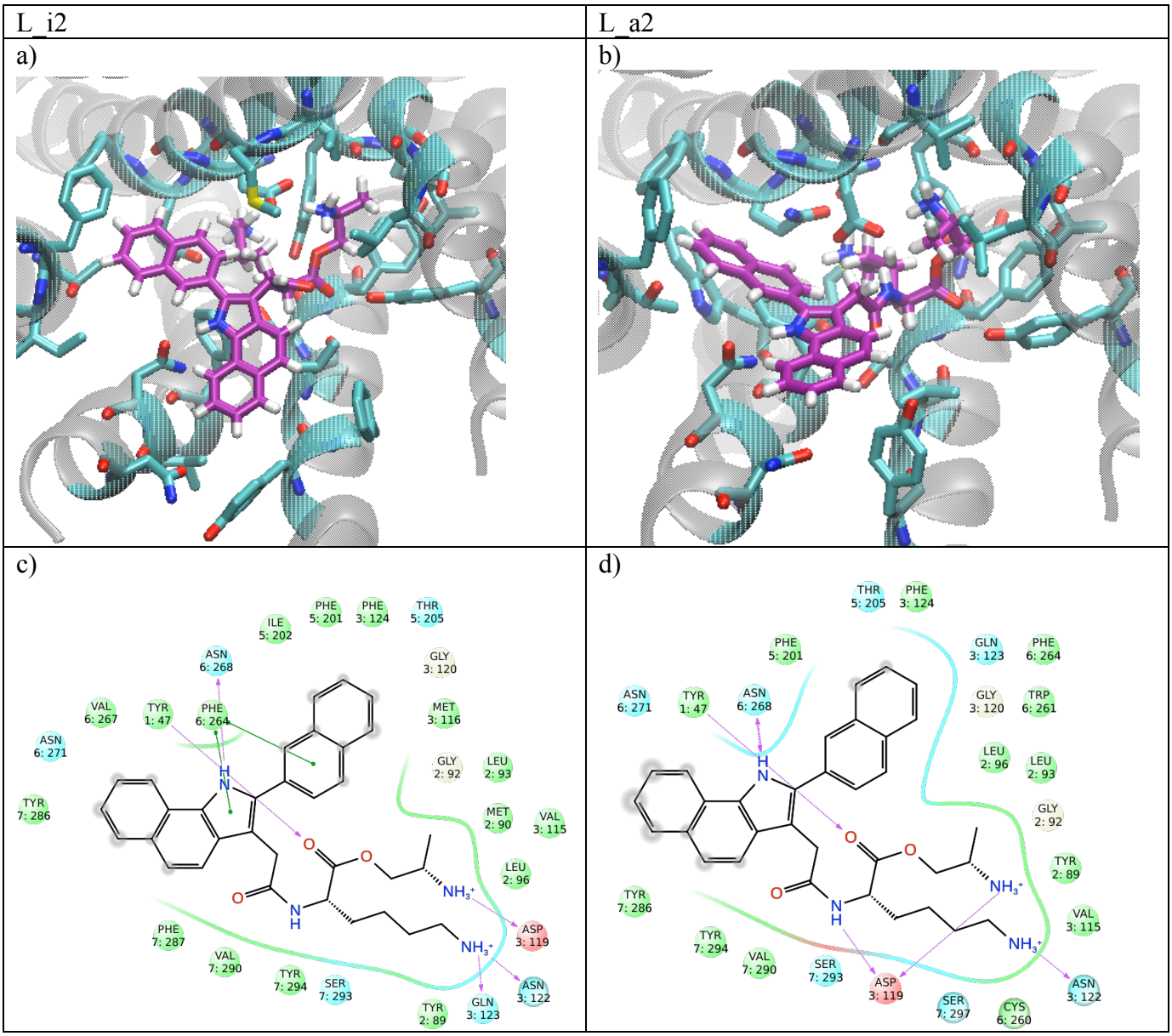

FIGURE S8 a) 3D visualization of the docking pose L_i2, the lowest energy complex with agonist L-817,818. b) 3D visualization of the docking pose L_a2. The ligand is shown in purple. c) Ligand interaction diagram (LID) of the pose L_i2. d) LID of the pose L_a2. LIDs are generated by Maestro9.3.(26) The cutoff distance for the residues shown is 4.0 Å. Hydrophobic interaction: green; polar interaction: blue; hydrogen bonds (cutoff distance 2.5 Å): purple arrows; π-π stacking: straight green lines.

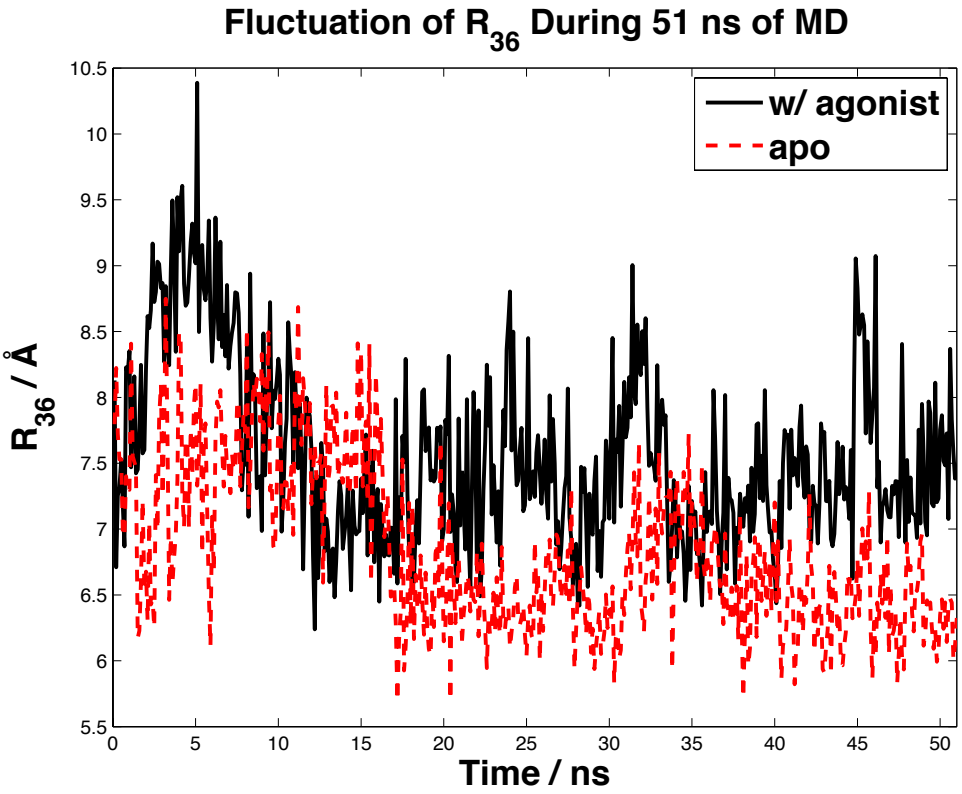

FIGURE S9 Fluctuation of R_{36} of hSSTR5 during 51 ns molecular dynamics simulation starting from the predicted inactive-state structure.

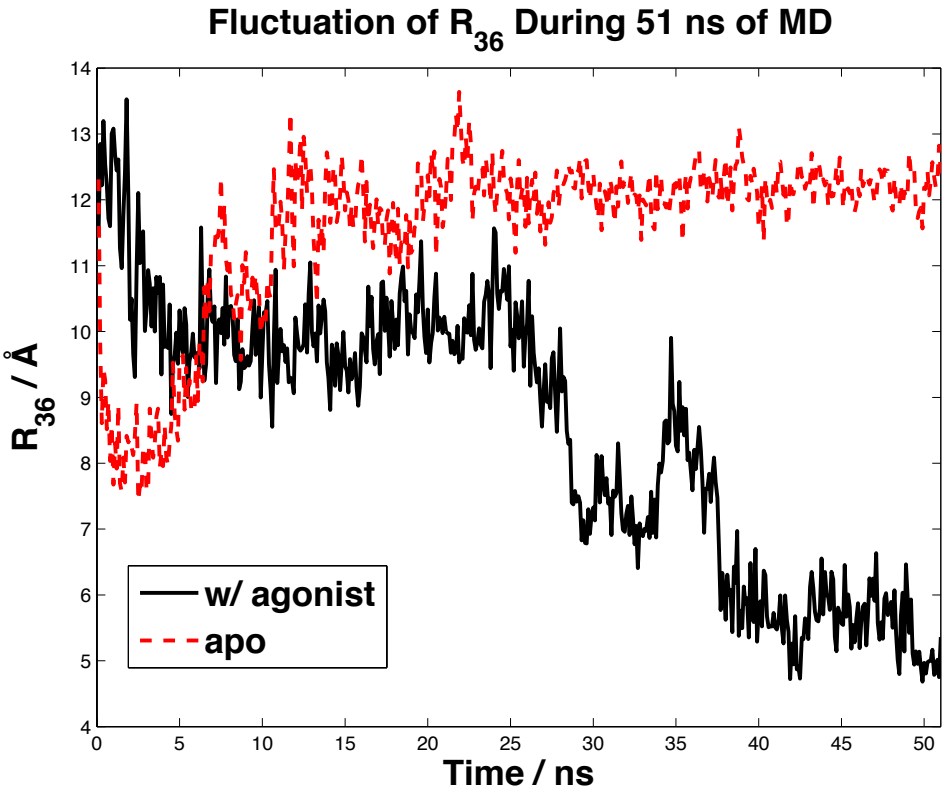

FIGURE S10 Fluctuation of R_{36} of hSSTR5 during 51 ns molecular dynamics simulation starting from the predicted active-state structure without the presence of the Gα protein.

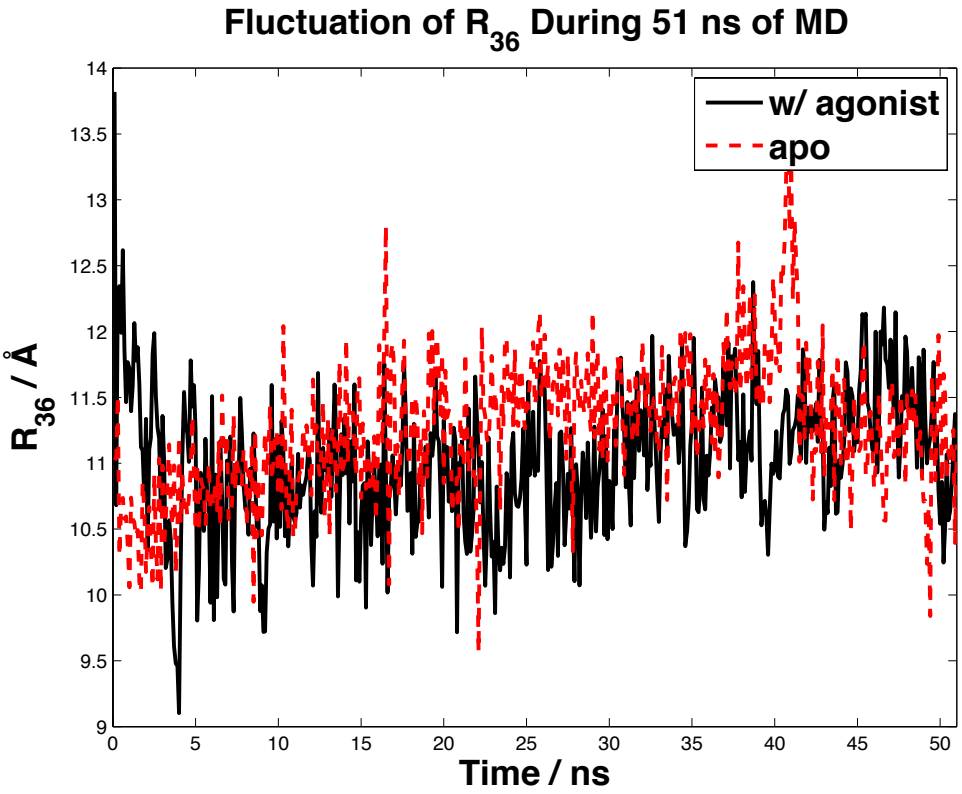

FIGURE S11 Fluctuation of R_{36} of hSSTR5 during 51 ns molecular dynamics simulation starting from the predicted active-state structure in complex with the Gα protein.

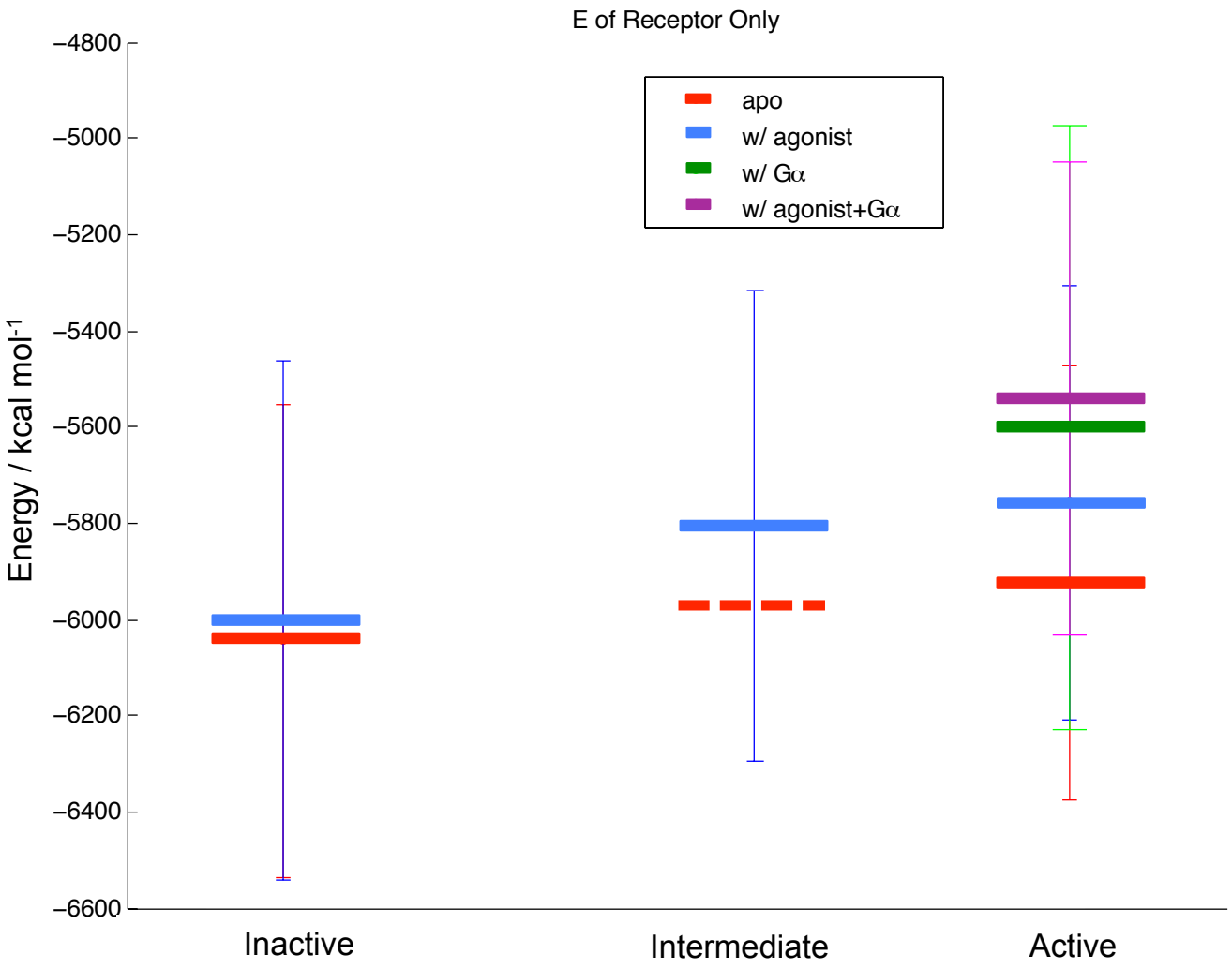

FIGURE S12 Energy profile of hSSTR5 during activation. The horizontal bars are E_R obtained according to the Materials and Methods section. The dashed horizontal bar is fictitious. "Inactive", "intermediate" and "active" states in the figure are defined as $R_{36} < 8$ Å, 8 Å $< R_{36} < 11$ Å, and $R_{36} > 11$ Å respectively.

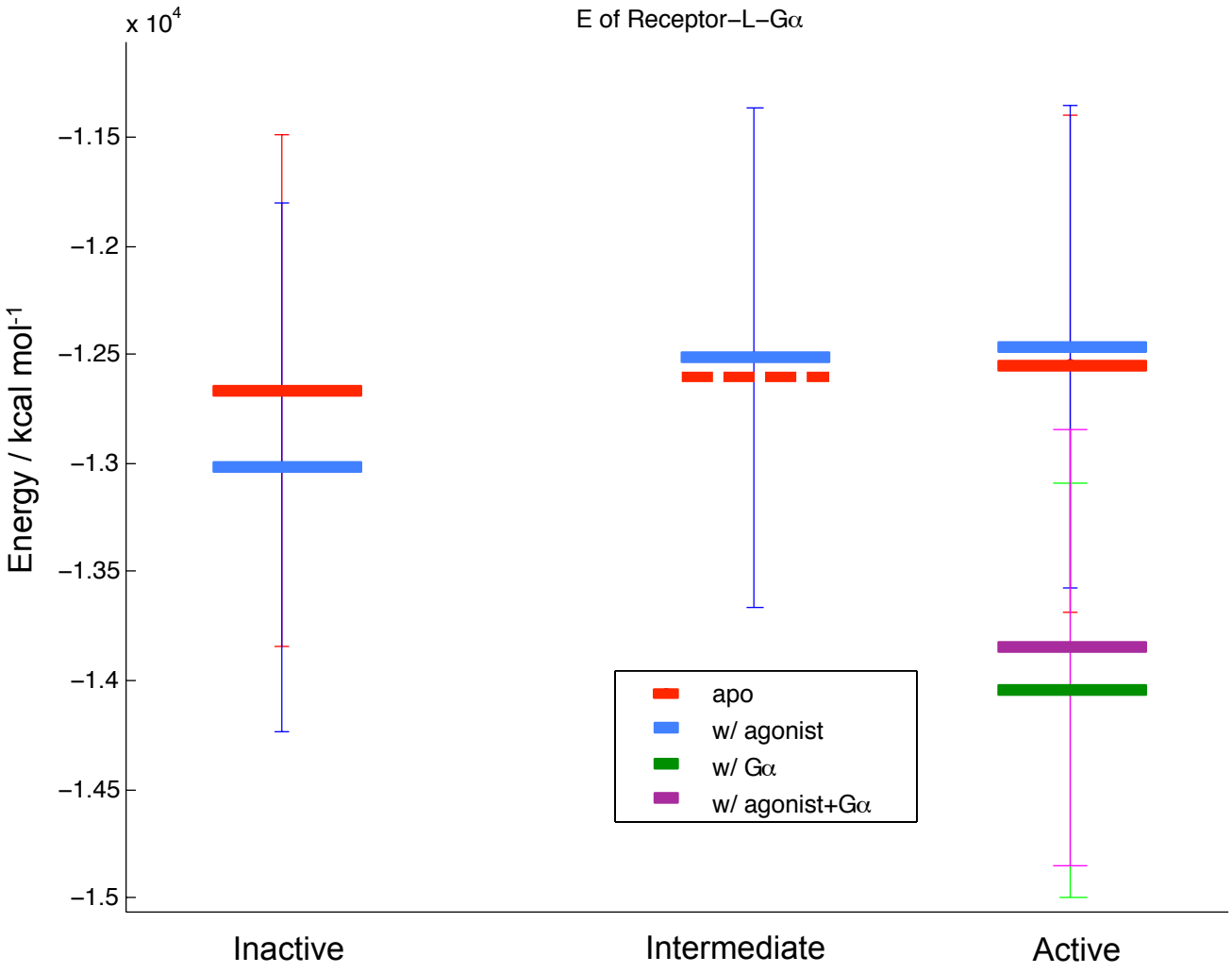

FIGURE S13 Energy profile of the system during hSSTR5 activation. The horizontal bars are E_{Total} calculated according to the Materials and Methods section. The dashed horizontal bar is fictitious. "Inactive", "intermediate" and "active" states in the figure are defined as $R_{36} < 8$ Å, 8 Å $< R_{36} < 11$ Å, and $R_{36} > 11$ Å respectively.

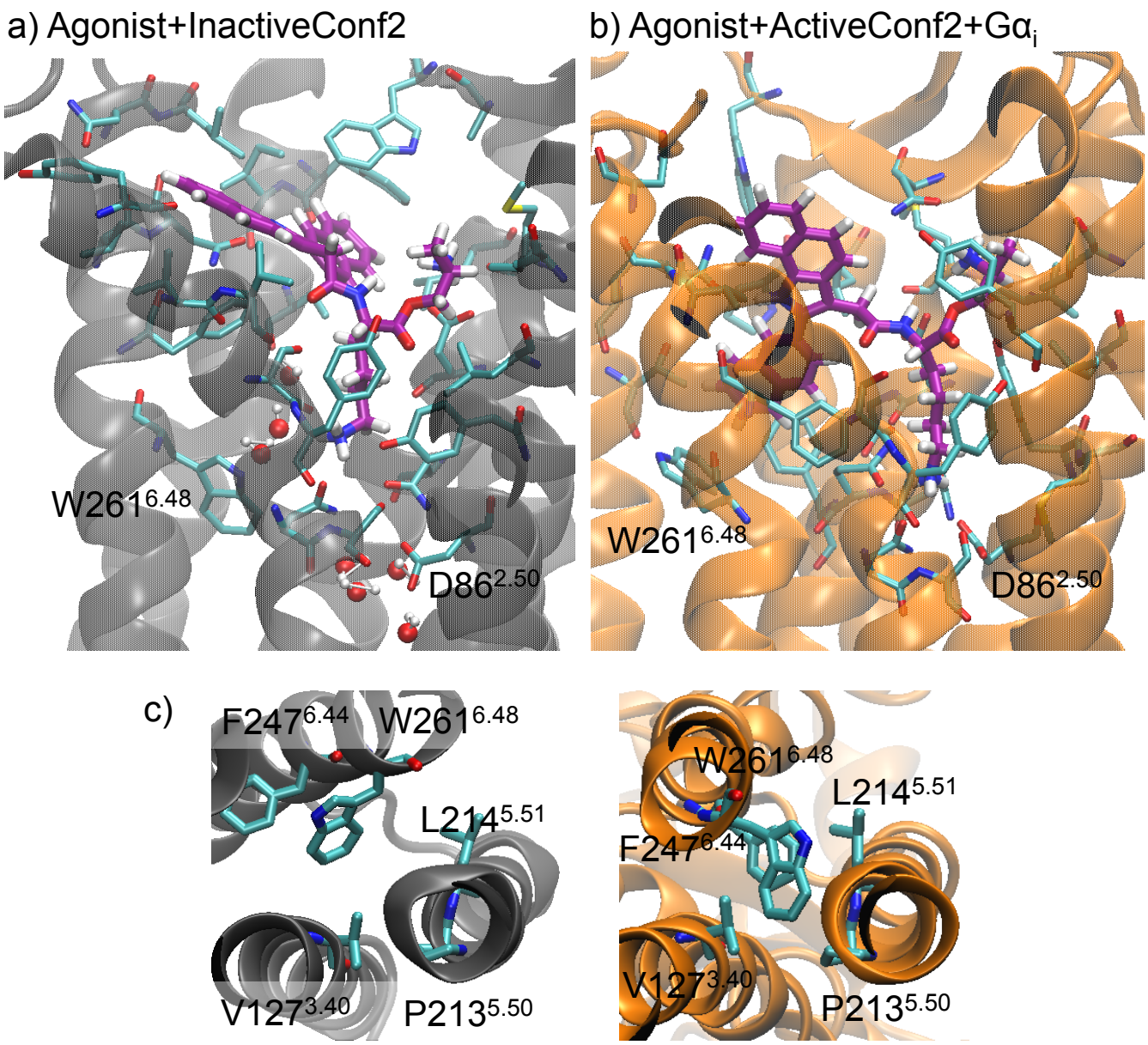

FIGURE S14 a) The salt bridge between L-817,818 (the agonist) and D86$^{2.50}$ of hSSTR5 is not present during the MD simulation of the agonist+InactiveConf2 complex. Instead, water molecules are surrounding D86$^{2.50}$. Water molecules within 10 Å of the side chain of D86$^{2.50}$ are displayed. b) A salt bridge between L-817,818 and D86$^{2.50}$ of hSSTR5 is formed during the MD simulation of the agonist+ActiveConf2+Gα$_i$ complex. In addition, there is π-π stacking between L-817,818 and W261$^{6.48}$ in the agonist+ActiveConf2+Gα$_i$ complex. There is no water molecule within 10 Å of the side chain of D86$^{2.50}$. c) The presence of the transmission switch: W261$^{6.48}$ and F257$^{6.44}$ are oriented towards P213$^{5.50}$ in ActiveConf2 (right panel) but not in InactiveConf2 (left panel) partly due to the rotation of TM6. The secondary structure in agonist+InactiveConf2 is shown in grey, and that in agonist+ActiveConf2+Gα$_i$ is shown in orange. Carbons in the residues on hSSTR5 are shown in cyan. The agonist carbon atoms are shown in purple.

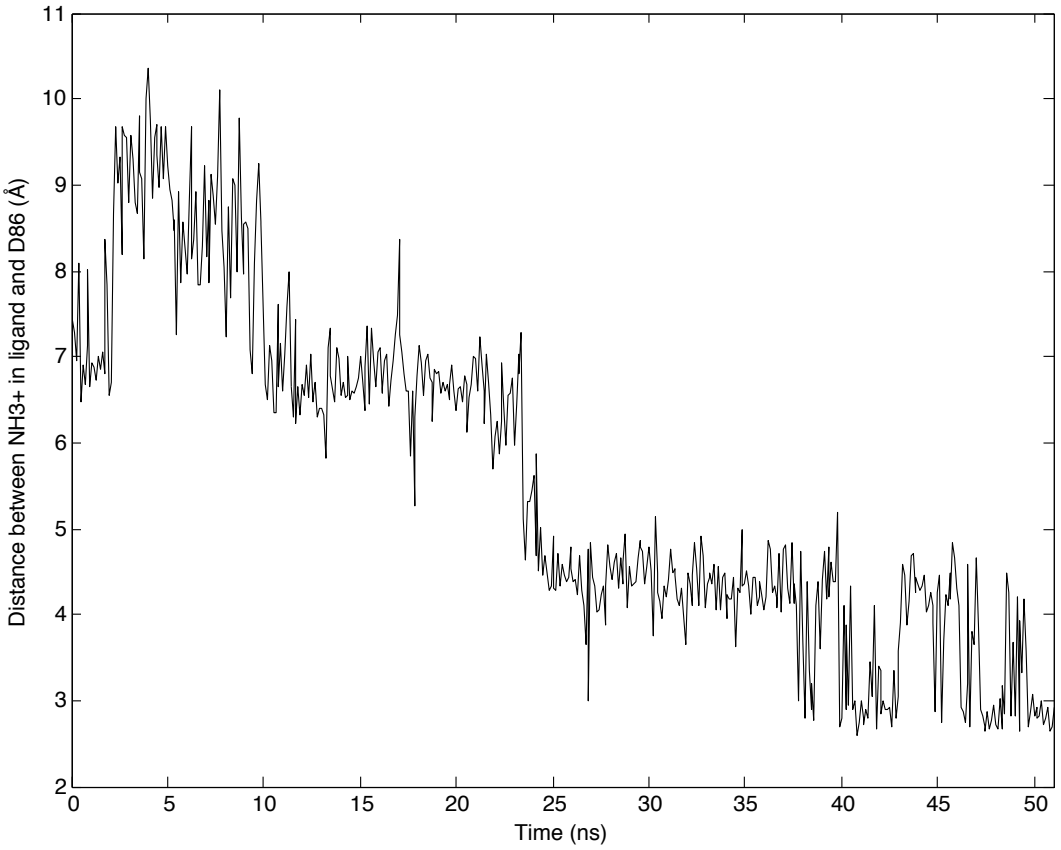

FIGURE S15 Distance between the N atom in an amine group of L-817,818 and a carboxylic acid oxygen atom in $D86^{2.50}$ of hSSTR5 along the trajectory of MD simulation of the agonist+ActiveConf2+$G\alpha_i$ complex.

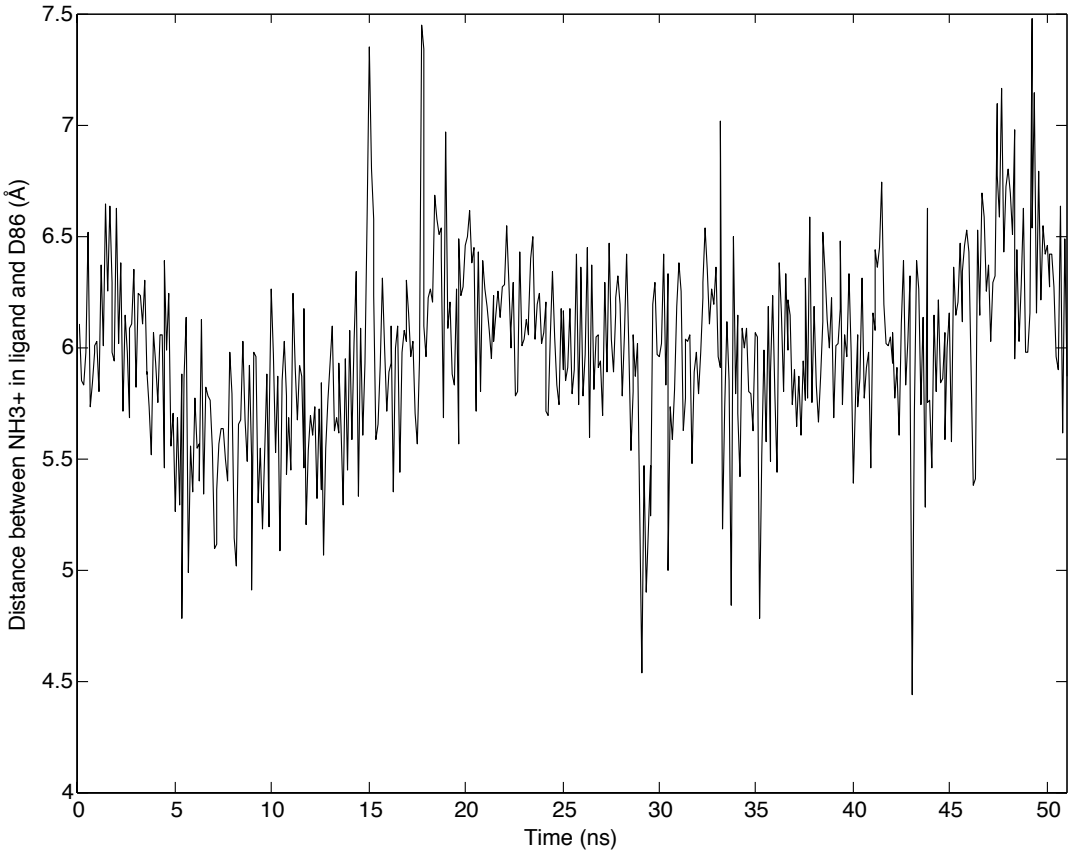

FIGURE S16 Distance between the N atom in an amine group of L-817,818 (the same amine group in Fig. S15) and a carboxylic acid oxygen atom in $D86^{2.50}$ of hSSTR5 along the trajectory of MD simulation of the agonist+InactiveConf2 complex.

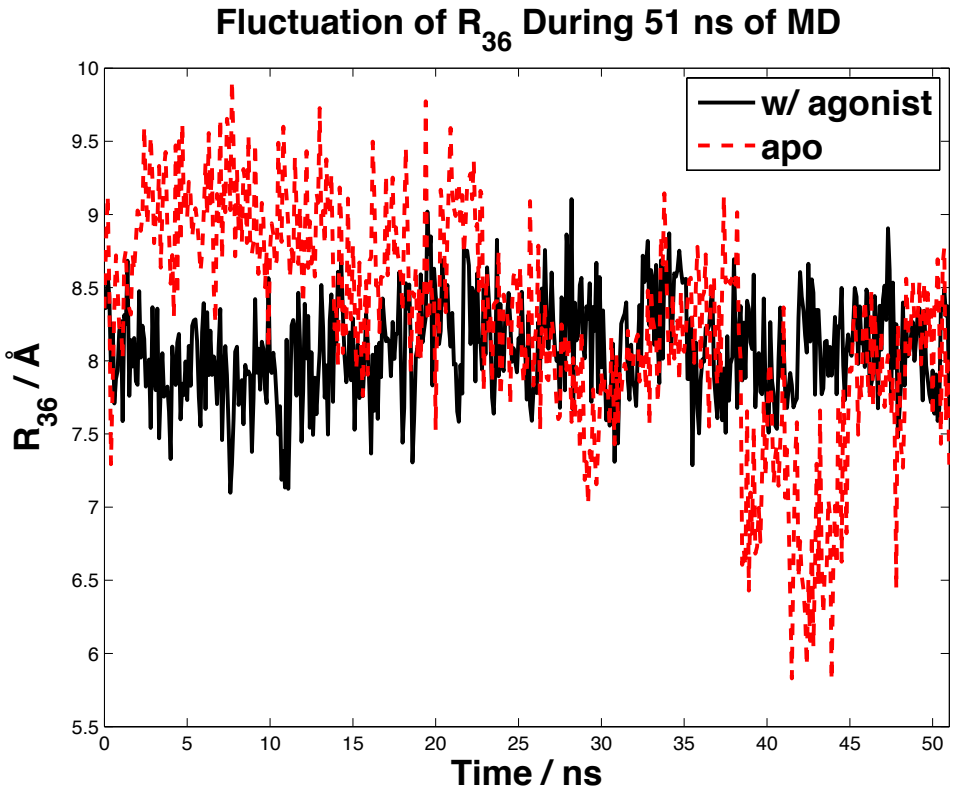

FIGURE S17 Fluctuation of R_{36} of hβ_2AR during 51 ns molecular dynamics simulation starting from the inactive-state structure.

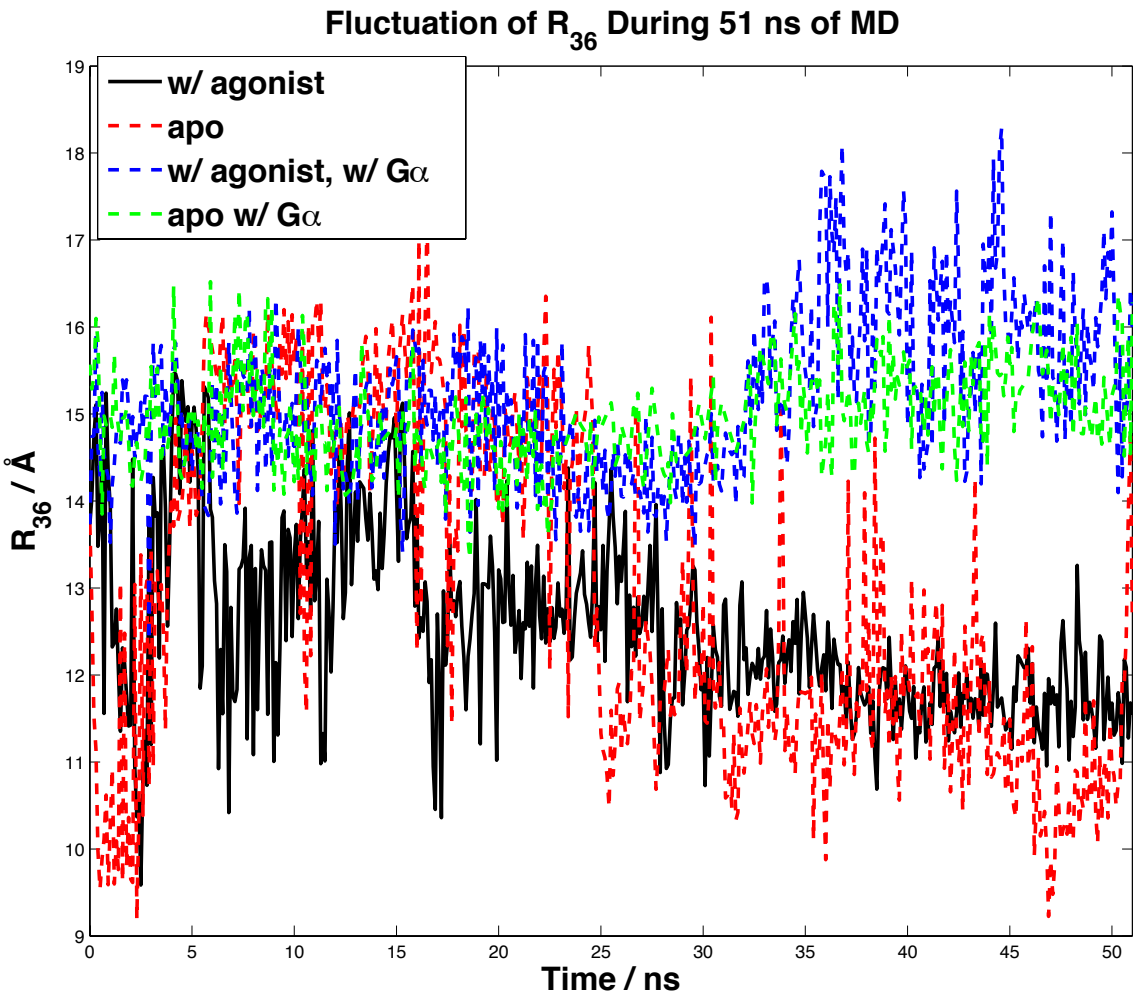

FIGURE S18 Fluctuation of R_{36} of hβ_2AR during 51 ns molecular dynamics simulation starting from the active-state structure.

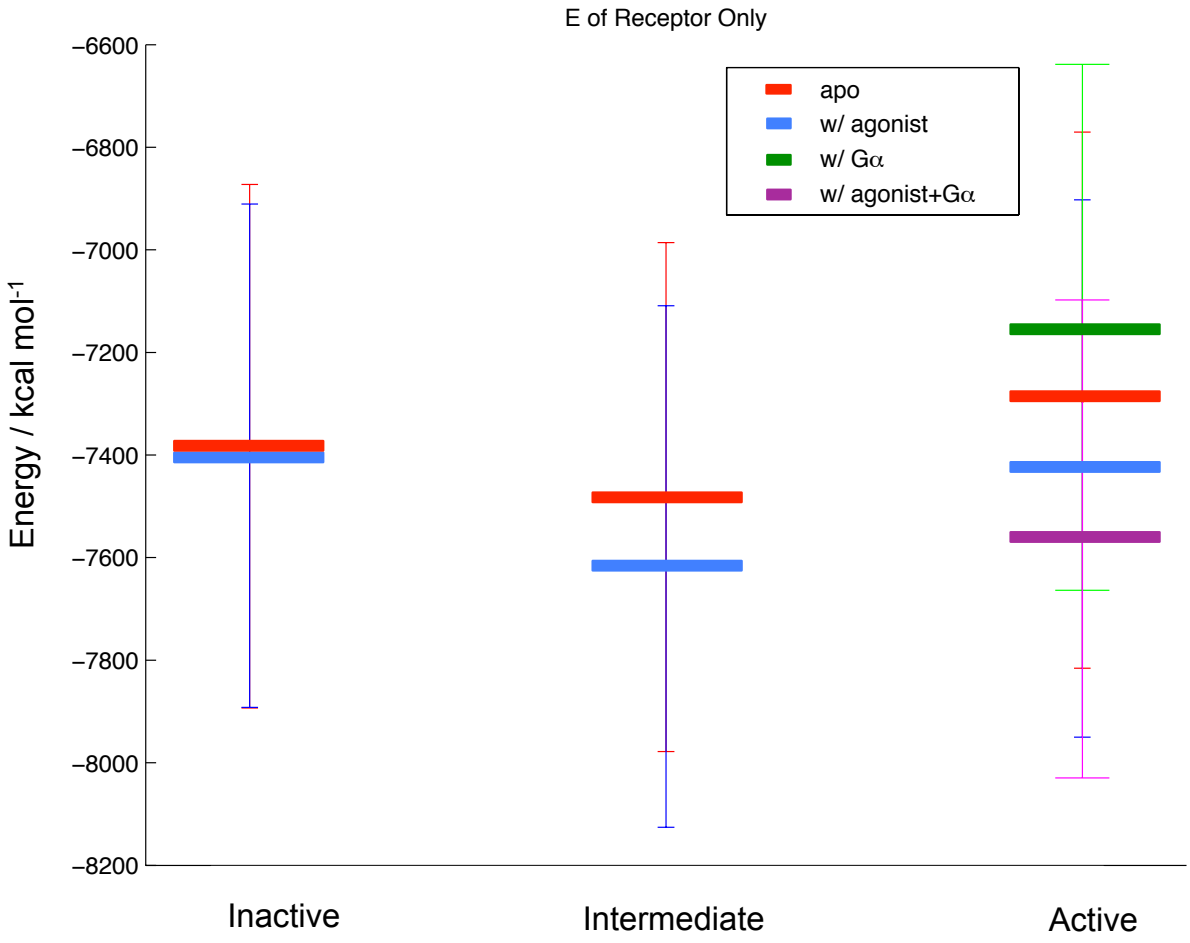

FIGURE S19 Energy profile of hβ$_2$AR during activation. The horizontal bars are E_R obtained according to the Materials and Methods section. "Inactive", "intermediate" and "active" states in the figure are defined as $R_{36} < 10$ Å, 10 Å $< R_{36} < 13$ Å, and $R_{36} > 13$ Å respectively.

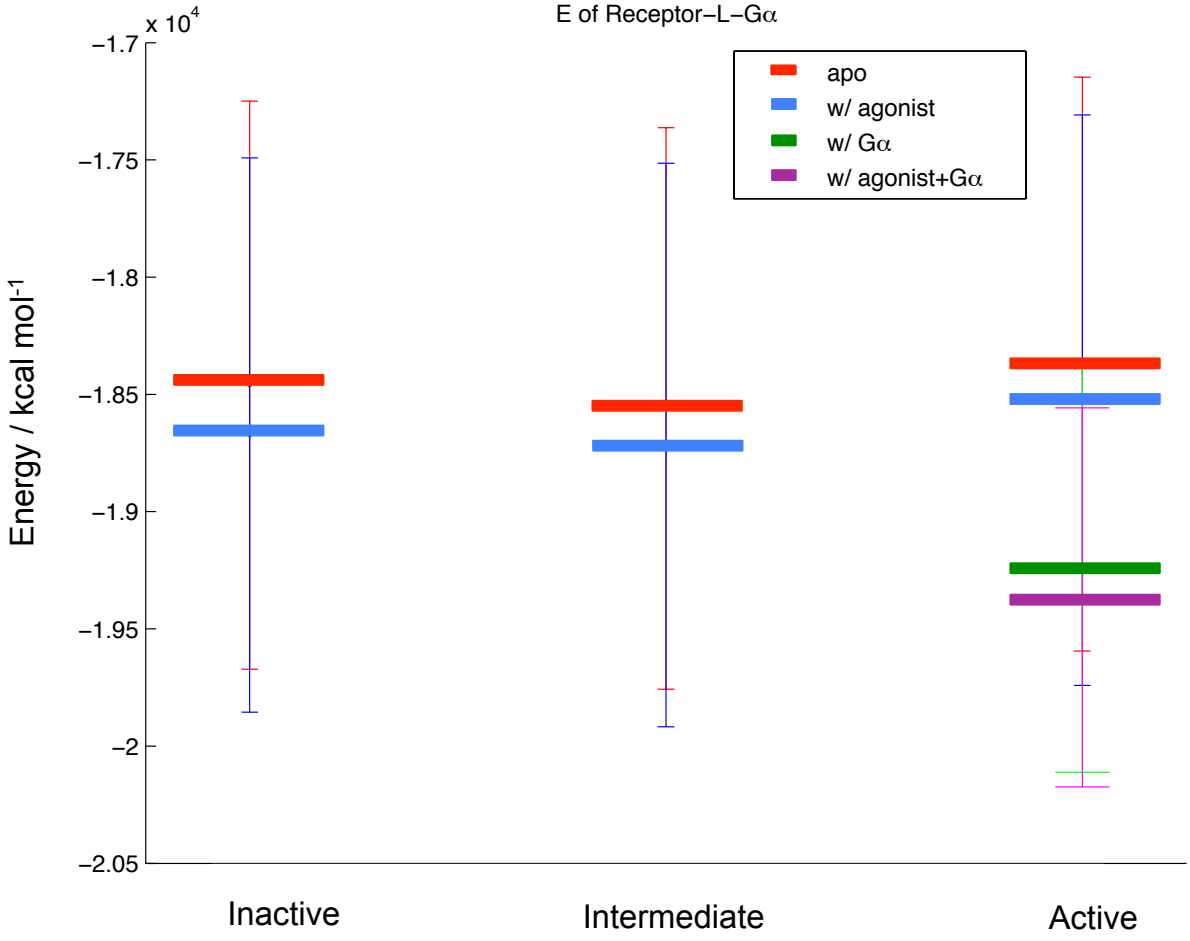

FIGURE S20 Energy profile of the system during hβ$_2$AR activation. The horizontal bars are E_{Total} calculated according to the Materials and Methods section. "Inactive", "intermediate" and "active" states in the figure are defined as $R_{36} < 10$ Å, 10 Å $< R_{36} < 13$ Å, and $R_{36} > 13$ Å respectively.

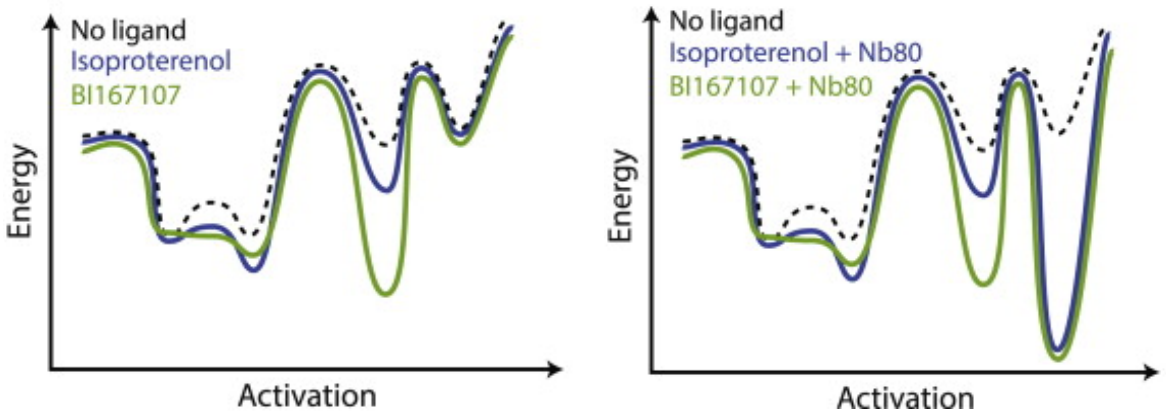

FIGURE S21 Qualitative energy landscape of hβ_2AR from experiments. This figure is adapted from Manglik et al. 2015. Cell 161:1101-1111.(27)

SUPPORTING REFERENCES

1. Lomize, M. A., A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg. 2006. OPM: orientations of proteins in membranes database. Bioinformatics 22:623-625.
2. Case, D. A., T. Darden, T. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. Walker, W. Zhang, and K. Merz. 2010. Amber 11. University of California.
3. Lim, K.-T., S. Brunett, M. Iotov, R. B. McClurg, N. Vaidehi, S. Dasgupta, S. Taylor, and W. A. Goddard. 1997. Molecular dynamics for very large systems on massively parallel computers: the MPSim program. J. Comput. Chem. 18:501-521.
4. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. J. Mol. Graphics 14:33-38.
5. Floriano, W. B., N. Vaidehi, G. Zamanakos, and W. A. Goddard. 2004. HierVLS hierarchical docking protocol for virtual ligand screening of large-molecule databases. J. Med. Chem. 47:56-71.
6. Goddard, W. A., S.-K. Kim, Y. Li, B. Trzaskowski, A. R. Griffith, and R. Abrol. 2010. Predicted 3D structures for adenosine receptors bound to ligands: Comparison to the crystal structure. J. Struct. Biol. 170:10-20.
7. MacroModel, version 9.7, Schrodinger, LLC, New York, NY, 2009.
8. Maestro, version 9.1, Schrodinger, LLC, New York, NY, 2010.
9. Chang, G., W. C. Guida, and W. C. Still. 1989. An internal-coordinate Monte Carlo method for searching conformational space. J. Am. Chem. Soc. 111:4379-4386.
10. Saunders, M., K. N. Houk, Y. D. Wu, W. C. Still, M. Lipton, G. Chang, and W. C. Guida. 1990. Conformations of cycloheptadecane. A comparison of methods for conformational searching. J. Am. Chem. Soc. 112:1419-1427.
11. Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. 118:11225-11236.
12. Jaguar, version 7.6, Schrodinger, LLC, New York, NY, 2009.
13. Schwede, T., J. Kopp, N. Guex, and M. C. Peitsch. 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31:3381-3385.
14. Manglik, A., A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka, and S. Granier. 2012. Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. Nature 485:321-326.
15. Rasmussen, S. G. F., B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. A. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara, and B. K. Kobilka. 2011. Crystal structure of the β_2 adrenergic receptor-Gs protein complex. Nature 477:549-555.
16. Cherezov, V., D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H.-J. Choi, P. Kuhn, W. I. Weis, and B. K. Kobilka. 2007. High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. Science 318:1258-1265.
17. Choe, H.-W., Y. J. Kim, J. H. Park, T. Morizumi, E. F. Pai, N. Krauß, K. P. Hofmann, P. Scheerer, and O. P. Ernst. 2011. Crystal structure of metarhodopsin II. Nature 471:651-655.
18. Okada, T., M. Sugihara, A.-N. Bondar, M. Elstner, P. Entel, and V. Buss. 2004. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. J. Mol. Biol. 342:571-583.

19. Kruse, A. C., A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hübner, E. Pardon, C. Valant, and P. M. Sexton. 2013. Activation and allosteric modulation of a muscarinic acetylcholine receptor. Nature 504:101-106.
20. Haga, K., A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, and T. Haga. 2012. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. Nature 482:547-551.
21. Huang, W., A. Manglik, A. Venkatakrishnan, T. Laeremans, E. N. Feinberg, A. L. Sanborn, H. E. Kato, K. E. Livingston, T. S. Thorsen, and R. C. Kling. 2015. Structural insights into μ-opioid receptor activation. Nature 524:315-321.
22. Lebon, G., T. Warne, P. C. Edwards, K. Bennett, C. J. Langmead, A. G. Leslie, and C. G. Tate. 2011. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. Nature 474:521-525.
23. Jaakola, V.-P., M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman, and R. C. Stevens. 2008. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 322:1211-1217.
24. Kang, Y., X. E. Zhou, X. Gao, Y. He, W. Liu, A. Ishchenko, A. Barty, T. A. White, O. Yefanov, and G. W. Han. 2015. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. Nature 523:561-567.
25. Dong, S. S., R. Abrol, and W. A. Goddard. 2015. The Predicted Ensemble of Low-Energy Conformations of Human Somatostatin Receptor Subtype 5 and the Binding of Antagonists. ChemMedChem 10:650-661.
26. Maestro, version 9.3, Schrödinger, LLC, New York, NY, 2012.
27. Manglik, A., T. H. Kim, M. Masureel, C. Altenbach, Z. Yang, D. Hilger, M. T. Lerch, T. S. Kobilka, F. S. Thian, and W. L. Hubbell. 2015. Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. Cell 161:1101-1111.